

Measuring the Effects of Starting Pitching

Lee Przybylski

November 14, 2021

Abstract

Betting on baseball is challenging. One feature that makes the sport different is that moneylines usually list probable starting pitchers. To take advantage of this, we develop a generalized linear mixed effects model using retrosheet data from several seasons. The model includes effects for teams, starting pitchers, and venue. Being able to assess a pitcher's performance independent of his team is also challenging. By estimating effects for each starting pitcher, fitting the model provides another way measure a starting pitcher's effectiveness. We also provide some background on pitching metrics that have been used in the past, such as ERA, FIP, and oppent WOBA, and compare these metrics to our estimated pitcher effects.

1 Introduction

Most of our data will be taken from [sportsbookreviewsonline.com](#)

2 Model Selection

2.1 Overdispersion

2.2 Predictive Value of the Model

Let y_{ijklm} be the number of runs scored by team i against team j at venue k facing starting pitcher l during the m th game of the season. The model we propose assumes that $y_{ijklm} \sim \text{Poisson}(\lambda_{ijklm})$ where

$$\log(\lambda_{ijklm}) = \mu + \tau_i + \delta_j + \rho_k + \chi_{im} + p_l + g_m,$$

$$p_l \stackrel{\text{iid}}{\sim} N(0, \sigma_p^2), \quad g_m \stackrel{\text{iid}}{\sim} N(0, \sigma_g^2).$$

3 Pitcher Effects

3.1 Noteworthy Pitchers

3.2 Comparison with Other Metrics